Package 'CARBayesST'

April 17, 2018

Type Package

Title Spatio-Temporal Generalised Linear Mixed Models for Areal Unit Data

Version 2.5.2

Date 2018-04-17

Author Duncan Lee, Alastair Rushworth and Gary Napier

Maintainer Duncan Lee <Duncan.Lee@glasgow.ac.uk>

Description

Implements a class of spatio-temporal generalised linear mixed models for areal unit data, with inference in a Bayesian setting using Markov chain Monte Carlo (MCMC) simulation. The response variable can be binomial, Gaussian or Poisson, but for some models only the binomial and Poisson data likelihoods are available. The spatio-temporal autocorrelation is modelled by random effects, which are assigned conditional autoregressive (CAR) style prior distributions. A number of different random effects structures are available, including Bernardinelli et al. (1995) <doi:10.1002/sim.4780142112>, Rushworth et al. (2014) <doi:10.1016/j.sste.2014.05.001> and Lee et al. (2016) <doi:10.1214/16-AOAS941>. Full details are given in the vignette accompanying this package. The creation of this package was supported by the Engineering and Physical Sciences Research Council (EPSRC) grant EP/J017442/1 and the Medical Research Council (MRC) grant MR/L022184/1.

License GPL (>= 2)

Depends MASS, R (>= 3.0.0), Rcpp (>= 0.11.5)

Imports CARBayesdata, coda, dplyr, matrixcalc, sp, spam, spdep, stats, testthat, truncdist, truncnorm, utils

LinkingTo Rcpp

LazyLoad yes

ByteCompile yes

URL http://github.com/duncanplee/CARBayesST

BugReports http://github.com/duncanplee/CARBayesST/issues

NeedsCompilation yes

Repository CRAN

Date/Publication 2018-04-17 09:50:33 UTC

R topics documented:

CARBayesST-package	*Spatio-Temporal Generalised Linear Mixed Models For Areal Unit Data*

Description

Implements a class of spatio-temporal generalised linear mixed models for areal unit data, with inference in a Bayesian setting using Markov chain Monte Carlo (MCMC) simulation. The response variable can be binomial, Gaussian or Poisson, but for some models only the binomial and Poisson data likelihoods are available. The spatio-temporal autocorrelation is modelled by random effects, which are assigned conditional autoregressive (CAR) style prior distributions. A number of different random effects structures are available, and full details are given in the vignette accompanying this package and the references below. The creation of this package was supported by the Engineering and Physical Sciences Research Council (EPSRC) grant EP/J017442/1 and the Medical Research Council (MRC) grant MR/L022184/1.

Details

Package:	CARBayesST
Type:	Package
Version:	2.5.2
Date:	2018-04-17
License:	GPL (>= 2)

Author(s)

Author: Duncan Lee, Alastair Rushworth and Gary Napier

Maintainer: Duncan Lee <Duncan.Lee@glasgow.ac.uk>

References

Bernardinelli, L., D. Clayton, C.Pascuto, C.Montomoli, M.Ghislandi, and M. Songini (1995). Bayesian analysis of space-time variation in disease risk. Statistics in Medicine, 14, 2433-2443.

Knorr-Held, L. (2000). Bayesian modelling of inseparable space-time variation in disease risk. Statistics in Medicine, 19, 2555-2567.

Lee, D and Lawson, C (2016). Quantifying the spatial inequality and temporal trends in maternal smoking rates in Glasgow, Annals of Applied Statistics, 10, 1427-1446.

Lee, D and Rushworth, A and Napier, G (2018). Spatio-Temporal Areal Unit Modeling in R with Conditional Autoregressive Priors Using the CARBayesST Package, Journal of Statistical Software, 84, 9, 1-39.

Napier, G, D. Lee, C. Robertson, A. Lawson, and K. Pollock (2016). A model to estimate the impact of changes in MMR vaccination uptake on inequalities in measles susceptibility in Scotland, Statistical Methods in Medical Research, 25, 1185-1200.

Roberts, G and Rosenthal, J (1998). Optimal scaling of discrete approximations to the Langevin diffusions, Journal of the Royal Statistical Society Series B 60, 255-268.

Rushworth, A., D. Lee, and R. Mitchell (2014). A spatio-temporal model for estimating the long-term effects of air pollution on respiratory hospital admissions in Greater London. Spatial and Spatio-temporal Epidemiology 10, 29-38.

Rushworth, A., Lee, D., and Sarran, C (2017). An adaptive spatio-temporal smoothing model for estimating trends and step changes in disease risk. Journal of the Royal Statistical Society Series C, 66, 141-157.

Examples

```
## See the examples in the function specific help files and in the vignette
## accompanying this package.
```

| coef.CARBayesST | *Extract the regression coefficients from a model.* |

Description

This function takes a CARBayesST object and returns the vector of estimated regression coefficients based on posterior medians.

Usage

```
    ## S3 method for class 'CARBayesST'
coef(object, ...)
```

Arguments

object	A CARBayesST fitted model object.
...	Ignored.

Author(s)

Duncan Lee

---

fitted.CARBayesST   *Extract the fitted values from a model.*

Description

This function takes a CARBayesST object and returns the vector of fitted values based on posterior medians.

Usage

```
    ## S3 method for class 'CARBayesST'
fitted(object, ...)
```

Arguments

object	A CARBayesST fitted model object.
...	Ignored.

Author(s)

Duncan Lee

---

logLik.CARBayesST   *Extract the estimated loglikelihood from a fitted model.*

Description

This function takes a CARBayesST object and returns the estimated loglikelihood based on posterior medians.

Usage

```
    ## S3 method for class 'CARBayesST'
logLik(object, ...)
```

Arguments

object	A CARBayesST fitted model object.
...	Ignored.

Author(s)

Duncan Lee

---

model.matrix.CARBayesST

Extract the Model (design) matrix from a model.

Description

This function takes a CARBayesST object and returns the design matrix.

Usage

```
     ## S3 method for class 'CARBayesST'
model.matrix(object, ...)
```

Arguments

object	A CARBayesST fitted model object.
...	Ignored.

Author(s)

Duncan Lee

---

print.CARBayesST   *Print a summary of the fitted model to the screen.*

Description

This function takes a CARBayesST object and returns a summary of the fitted model. The summary includes, for selected parameters, posterior medians and 95 percent credible intervals, the number of samples, the acceptance rate, the effective number of independent samples and the Geweke convergence diagnostic in the form of a Z-score.

Usage

```
## S3 method for class 'CARBayesST'
print(x, ...)
```

Arguments

x	A CARBayesST fitted model object.
...	Ignored.

Author(s)

Duncan Lee

---

residuals.CARBayesST   *Extract the residuals from a model.*

Description

This function takes a CARBayesST object and returns a set of residuals. The allowable types of residual are "response" (raw), "pearson", and "deviance". In each case the fitted values are based on posterior medians.

Usage

```
    ## S3 method for class 'CARBayesST'
residuals(object, type, ...)
```

Arguments

object	A CARBayesST fitted model object.
type	A text string and one of c("response", "pearson", "deviance"). If this argument is omitted the default is "deviance".
...	Ignored.

Author(s)

Duncan Lee

ST.CARadaptive   *Fit a spatio-temporal generalised linear mixed model to data with a spatio-temporal autoregressive process that has an adaptive autocorrelation stucture.*

Description

Fit a spatio-temporal generalised linear mixed model to areal unit data, where the response variable can be binomial, Gaussian or Poisson. The linear predictor is modelled by known covariates and a vector of random effects. The latter follows a multivariate first order autoregressive time series process, where spatial autocorrelation is modelled via the precision matrix which is based on an adjacency matrix W. The non-zero elements of W associated with each pair of geographically adjacent areal units are treated as random variables, which allows step changes to be identified in the random effects surface between geographically adjacent regions. The model is similar to that proposed by Rushworth et al. (2017). Further details are given in the vignette accompanying this package. Inference is conducted in a Bayesian setting using Markov chain Monte Carlo (MCMC) simulation.

Usage

```
ST.CARadaptive(formula, family, data=NULL, trials=NULL, W, burnin, n.sample, thin=1,
prior.mean.beta=NULL, prior.var.beta=NULL, prior.nu2=NULL, prior.tau2=NULL,
rhofix = NULL, epsilon = 0, MALA=TRUE, verbose=TRUE)
```

Arguments

formula
: A formula for the covariate part of the model using the syntax of the lm() function. Offsets can be included here using the offset() function. The response and each covariate should be vectors of length (KN)*1, where K is the number of spatial units and N is the number of time periods. Each vector should be ordered so that the first K data points are the set of all K spatial locations at time 1, the next K are the set of spatial locations for time 2 and so on.

family
: One of either 'binomial', 'gaussian' or 'poisson', which respectively specify a binomial likelihood model with a logistic link function, a Gaussian likelihood model with an identity link function, or a Poisson likelihood model with a log link function.

data
: An optional data.frame containing the variables in the formula.

trials
: A vector the same length and in the same order as the response containing the total number of trials for each area and time period. Only used if family='binomial'.

W
: A non-negative K by K neighbourhood matrix (where K is the number of spatial units). Typically a binary specification is used, where the jkth element equals one if areas (j, k) are spatially close (e.g. share a common border) and is zero otherwise. The matrix can be non-binary, but each row must contain at least one non-zero entry.

burnin
: The number of MCMC samples to discard as the burn-in period.

n.sample
: The number of MCMC samples to generate.

thin
: The level of thinning to apply to the MCMC samples to reduce their temporal autocorrelation. Defaults to 1 (no thinning).

prior.mean.beta
: A vector of prior means for the regression parameters beta (Gaussian priors are assumed). Defaults to a vector of zeros.

prior.var.beta	A vector of prior variances for the regression parameters beta (Gaussian priors are assumed). Defaults to a vector with values 1000.
prior.nu2	The prior shape and scale in the form of c(shape, scale) for an Inverse-Gamma(shape, scale) prior for the Gaussian error variance nu2. Defaults to c(1, 0.01) and only used if family='Gaussian'.
prior.tau2	The prior shape and scale in the form of c(shape, scale) for an Inverse-Gamma(shape, scale) prior for the random effect variance tau2. Defaults to c(1, 0.01).
rhofix	Value between 0 and 1 at which to fix the spatial correlation parameter. Setting rhofix=1, reduces the random effects prior to the intrinsic CAR model but does require epsilon>0. When set to NULL, (the default) the parameter is estimated with a uniform[0,1] prior.
epsilon	Diagonal ridge parameter to add to the random effects prior precision matrix, only required when rhofix = 1, and the prior precision is improper. Defaults to 0.
MALA	Logical, should the function use MALA (TRUE, default) or simple random walk (FALSE) updates for the random effects. Not applicable if family='gaussian'.
verbose	Logical, should the function update the user on its progress.

Value

summary.results	
	A summary table of the parameters.
samples	A list containing the MCMC samples from the model.
fitted.values	A vector of fitted values for each area and time period.
residuals	A matrix with 3 columns where each column is a type of residuals for each area and time period. The types are "response" (raw), "pearson", and "deviance".
modelfit	Model fit criteria including the Deviance Information Criterion (DIC) and its corresponding estimated effective number of parameters (p.d), the Log Marginal Predictive Likelihood (LMPL), the Watanabe-Akaike Information Criterion (WAIC) and its corresponding estimated number of effective parameters (p.w), and the estimated loglikelihood.
localised.structure	
	A list with 2 K*K matrices, Wmedian and W99 summarising the estimated adjacency relationships. Wmedian contains the posterior median for each w_ij element estimated in the model for adjacent areal units, while W99 contains binary indicator variables for whether Prob(w_ij < 0.5|data)>0.99. For both matrices, elements corresponding to non-adjacent pairs of areas have NA values.
formula	The formula for the covariate and offset part of the model.
model	A text string describing the model fit.
X	The design matrix of covariates.

Author(s)

Alastair Rushworth

References

Rushworth, A., Lee, D., and Sarran, C (2017). An adaptive spatio-temporal smoothing model for estimating trends and step changes in disease risk. Journal of the Royal Statistical Society Series C, 66, 141-157.

Examples

```
##################################################
#### Run the model on simulated data on a lattice
##################################################
#### set up the regular lattice
x.easting <- 1:10
x.northing <- 1:10
Grid <- expand.grid(x.easting, x.northing)
K <- nrow(Grid)
N <- 10
N.all <- N * K


#### set up spatial neighbourhood matrix W
distance <- as.matrix(dist(Grid))
W <-array(0, c(K,K))
W[distance==1] <-1


#### Simulate the elements in the linear predictor and the data
Q.W <- 0.99 * (diag(apply(W, 2, sum)) - W) + 0.01 * diag(rep(1,K))
Q.W.inv <- solve(Q.W)
phi.temp <- mvrnorm(n=1, mu=rep(0,K), Sigma=(0.1 * Q.W.inv))
phi <- phi.temp
    for(i in 2:N)
    {
    phi.temp2 <- mvrnorm(n=1, mu=(0.8 * phi.temp), Sigma=(0.1 * Q.W.inv))
    phi.temp <- phi.temp2
    phi <- c(phi, phi.temp)
    }
jump <- rep(c(rep(2, 70), rep(4, 30)),N)
LP <- jump + phi
fitted <- exp(LP)
Y <- rpois(n=N.all, lambda=fitted)


#### Run the model
## Not run: model <- ST.CARadaptive(formula=Y~1, family="poisson", W=W, burnin=10000,
n.sample=50000)
## End(Not run)


#### Toy example for checking
model <- ST.CARadaptive(formula=Y~1, family="poisson", W=W, burnin=10,
n.sample=50)
```

ST.CARanova	*Fit a spatio-temporal generalised linear mixed model to data, with spatial and temporal main effects and a spatio-temporal interaction.*

Description

Fit a spatio-temporal generalised linear mixed model to areal unit data, where the response variable can be binomial, Gaussian or Poisson. The linear predictor is modelled by known covariates and three vectors of random effects. The latter include spatial and temporal main effects and a spatio-temporal interaction. The spatial and temporal main effects are modelled by the conditional autoregressive (CAR) prior proposed by Leroux et al. (2000), while the spatio-temporal interaction random effects are independent. Due to the lack of identifiability between the interactions and the Gaussian errors, only main effects are allowed in the Gaussian model. Missing values are allowed in the response in this model. The model is similar to that proposed by Knorr-Held (2000) and further details are given in the vignette accompanying this package. Inference is conducted in a Bayesian setting using Markov chain Monte Carlo (MCMC) simulation.

Usage

```
ST.CARanova(formula, family, data=NULL,  trials=NULL, W, interaction=TRUE, burnin,
            n.sample, thin=1, prior.mean.beta=NULL, prior.var.beta=NULL,
            prior.nu2=NULL, prior.tau2=NULL, fix.rho.S=FALSE, rho.S=NULL,
            fix.rho.T=FALSE, rho.T=NULL, MALA=TRUE, verbose=TRUE)
```

Arguments

formula	A formula for the covariate part of the model using the syntax of the lm() function. Offsets can be included here using the offset() function. The response and each covariate should be vectors of length (KN)*1, where K is the number of spatial units and N is the number of time periods. Each vector should be ordered so that the first K data points are the set of all K spatial locations at time 1, the next K are the set of spatial locations for time 2 and so on. Missing values are allowed in the response.
family	One of either 'binomial', 'gaussian' or 'poisson', which respectively specify a binomial likelihood model with a logistic link function, a Gaussian likelihood model with an identity link function, or a Poisson likelihood model with a log link function.
data	An optional data.frame containing the variables in the formula.
trials	A vector the same length and in the same order as the response containing the total number of trials for each area and time period. Only used if family='binomial'.
W	A non-negative K by K neighbourhood matrix (where K is the number of spatial units). Typically a binary specification is used, where the jkth element equals one if areas (j, k) are spatially close (e.g. share a common border) and is zero otherwise. The matrix can be non-binary, but each row must contain at least one non-zero entry. Note, if fix.rho.S=TRUE and rho.S=0 then the W matrix is not

	actually used by the model so a K by K matrix of zeros could be entered instead if a real W matrix cannot be obtained.
interaction	TRUE or FALSE indicating whether the spatio-temporal interaction random effects should be included. Defaults to TRUE unless family='gaussian' in which case interactions are not allowed.
burnin	The number of MCMC samples to discard as the burn-in period.
n.sample	The number of MCMC samples to generate.
thin	The level of thinning to apply to the MCMC samples to reduce their temporal autocorrelation. Defaults to 1 (no thinning).
prior.mean.beta	
	A vector of prior means for the regression parameters beta (Gaussian priors are assumed). Defaults to a vector of zeros.
prior.var.beta	A vector of prior variances for the regression parameters beta (Gaussian priors are assumed). Defaults to a vector with values 1000.
prior.nu2	The prior shape and scale in the form of c(shape, scale) for an Inverse-Gamma(shape, scale) prior for the Gaussian error variance nu2. Defaults to c(1, 0.01) and only used if family='Gaussian'.
prior.tau2	The prior shape and scale in the form of c(shape, scale) for an Inverse-Gamma(shape, scale) prior for the random effect variances tau2. Defaults to c(1, 0.01).
fix.rho.S	Logical, should the spatial dependence parameter rho.S be fixed when fitting the model.
rho.S	If fix.rho.S=TRUE, then the value in the interval [0,1] that rho.S should be fixed to. Not needed if fix.rho.S=FALSE.
fix.rho.T	Logical, should the temporal dependence parameter rho.T be fixed when fitting the model.
rho.T	If fix.rho.T=TRUE, then the value in the interval [0,1] that rho.T should be fixed to. Not needed if fix.rho.T=FALSE.
MALA	Logical, should the function use MALA (TRUE, default) or simple random walk (FALSE) updates for the random effects. Not applicable if family='gaussian'.
verbose	Logical, should the function update the user on its progress.

Value

summary.results	
	A summary table of the parameters.
samples	A list containing the MCMC samples from the model.
fitted.values	A vector of fitted values for each area and time period.
residuals	A matrix with 3 columns where each column is a type of residuals for each area and time period. The types are "response" (raw), "pearson", and "deviance".
modelfit	Model fit criteria including the Deviance Information Criterion (DIC) and its corresponding estimated effective number of parameters (p.d), the Log Marginal Predictive Likelihood (LMPL), the Watanabe-Akaike Information Criterion (WAIC) and its corresponding estimated number of effective parameters (p.w), and the estimated loglikelihood.

accept	The acceptance probabilities for the parameters.
localised.structure	
	NULL, for compatability with the other models.
formula	The formula for the covariate and offset part of the model.
model	A text string describing the model fit.
X	The design matrix of covariates.

Author(s)

Duncan Lee

References

Knorr-Held, L. (2000). Bayesian modelling of inseparable space-time variation in disease risk. Statistics in Medicine, 19, 2555-2567.

Leroux, B., X. Lei, and N. Breslow (2000). Estimation of disease rates in small areas: A new mixed model for spatial dependence, Chapter Statistical Models in Epidemiology, the Environment and Clinical Trials, Halloran, M and Berry, D (eds), pp. 135-178. Springer-Verlag, New York.

Examples

```
##################################################
#### Run the model on simulated data on a lattice
##################################################
#### set up the regular lattice
x.easting <- 1:10
x.northing <- 1:10
Grid <- expand.grid(x.easting, x.northing)
K <- nrow(Grid)
N <- 10
N.all <- N * K


#### set up spatial (W) and temporal (D) neighbourhood matrices
distance <- as.matrix(dist(Grid))
W <-array(0, c(K,K))
W[distance==1] <-1

D <-array(0, c(N,N))
for(i in 1:N)
{
    for(j in 1:N)
    {
        if(abs((i-j))==1)  D[i,j] <- 1
    }
}


#### Simulate the elements in the linear predictor and the data
gamma <- rnorm(n=N.all, mean=0, sd=0.001)
```

```
x <- rnorm(n=N.all, mean=0, sd=1)
beta <- 0.1

Q.W <- 0.99 * (diag(apply(W, 2, sum)) - W) + 0.01 * diag(rep(1,K))
Q.W.inv <- solve(Q.W)
phi <- mvrnorm(n=1, mu=rep(0,K), Sigma=(0.01 * Q.W.inv))

Q.D <- 0.99 * (diag(apply(D, 2, sum)) - D) + 0.01 * diag(rep(1,N))
Q.D.inv <- solve(Q.D)
delta <- mvrnorm(n=1, mu=rep(0,N), Sigma=(0.01 * Q.D.inv))

phi.long <- rep(phi, N)
delta.long <- kronecker(delta, rep(1,K))
LP <- 4 +  x * beta + phi.long +  delta.long + gamma
mean <- exp(LP)
Y <- rpois(n=N.all, lambda=mean)


#### Run the model
## Not run: model <- ST.CARanova(formula=Y~x, family="poisson", interaction=TRUE,
W=W, burnin=10000, n.sample=50000, fix.rho.S=TRUE, rho.S=1)
## End(Not run)


#### Toy example for checking
model <- ST.CARanova(formula=Y~x, family="poisson", interaction=TRUE,
W=W, burnin=10, n.sample=50, fix.rho.S=TRUE, rho.S=1)
```

ST.CARar	*Fit a spatio-temporal generalised linear mixed model to data, with a spatio-temporal autoregressive process.*

Description

Fit a spatio-temporal generalised linear mixed model to areal unit data, where the response variable can be binomial, Gaussian or Poisson. The linear predictor is modelled by known covariates and a vector of random effects. The latter follows a multivariate first order autoregressive time series process, where spatial autocorrelation is modelled via the precision matrix. This precision matrix comes from the conditional autoregressive (CAR) prior proposed by Leroux et al. (2000), and the full model was proposed by Rushworth et al. (2014). Missing values are allowed in the response in this model. Further details are given in the vignette accompanying this package. Inference is conducted in a Bayesian setting using Markov chain Monte Carlo (MCMC) simulation.

Usage

```
ST.CARar(formula, family, data=NULL,  trials=NULL, W, burnin, n.sample, thin=1,
 prior.mean.beta=NULL, prior.var.beta=NULL, prior.nu2=NULL, prior.tau2=NULL,
 fix.rho.S=FALSE, rho.S=NULL, fix.rho.T=FALSE, rho.T=NULL, MALA=TRUE, verbose=TRUE)
```

Arguments

formula	A formula for the covariate part of the model using the syntax of the lm() function. Offsets can be included here using the offset() function. The response and each covariate should be vectors of length (KN)*1, where K is the number of spatial units and N is the number of time periods. Each vector should be ordered so that the first K data points are the set of all K spatial locations at time 1, the next K are the set of spatial locations for time 2 and so on. Missing values are allowed in the response.
family	One of either 'binomial', 'gaussian' or 'poisson', which respectively specify a binomial likelihood model with a logistic link function, a Gaussian likelihood model with an identity link function, or a Poisson likelihood model with a log link function.
data	An optional data.frame containing the variables in the formula.
trials	A vector the same length and in the same order as the response containing the total number of trials for each area and time period. Only used if family='binomial'.
W	A non-negative K by K neighbourhood matrix (where K is the number of spatial units). Typically a binary specification is used, where the jkth element equals one if areas (j, k) are spatially close (e.g. share a common border) and is zero otherwise. The matrix can be non-binary, but each row must contain at least one non-zero entry. Note, if fix.rho.S=TRUE and rho.S=0 then the W matrix is not actually used by the model so a K by K matrix of zeros could be entered instead if a real W matrix cannot be obtained.
burnin	The number of MCMC samples to discard as the burn-in period.
n.sample	The number of MCMC samples to generate.
thin	The level of thinning to apply to the MCMC samples to reduce their temporal autocorrelation. Defaults to 1 (no thinning).
prior.mean.beta	A vector of prior means for the regression parameters beta (Gaussian priors are assumed). Defaults to a vector of zeros.
prior.var.beta	A vector of prior variances for the regression parameters beta (Gaussian priors are assumed). Defaults to a vector with values 1000.
prior.nu2	The prior shape and scale in the form of c(shape, scale) for an Inverse-Gamma(shape, scale) prior for the Gaussian error variance nu2. Defaults to c(1, 0.01) and only used if family='Gaussian'.
prior.tau2	The prior shape and scale in the form of c(shape, scale) for an Inverse-Gamma(shape, scale) prior for the random effect variance tau2. Defaults to c(1, 0.01).
fix.rho.S	Logical, should the spatial dependence parameter rho.S be fixed when fitting the model.
rho.S	If fix.rho.S=TRUE, then the value in the interval [0,1] that rho.S should be fixed to. Not needed if fix.rho.S=FALSE.
fix.rho.T	Logical, should the temporal dependence parameter rho.T be fixed when fitting the model.

rho.T	If fix.rho.T=TRUE, then the value in the interval [0,1] that rho.T should be fixed to. Not needed if fix.rho.T=FALSE.
MALA	Logical, should the function use MALA (TRUE, default) or simple random walk (FALSE) updates for the random effects. Not applicable if family='gaussian'.
verbose	Logical, should the function update the user on its progress.

Value

summary.results	A summary table of the parameters.
samples	A list containing the MCMC samples from the model.
fitted.values	A vector of fitted values for each area and time period.
residuals	A matrix with 3 columns where each column is a type of residuals for each area and time period. The types are "response" (raw), "pearson", and "deviance".
modelfit	Model fit criteria including the Deviance Information Criterion (DIC) and its corresponding estimated effective number of parameters (p.d), the Log Marginal Predictive Likelihood (LMPL), the Watanabe-Akaike Information Criterion (WAIC) and its corresponding estimated number of effective parameters (p.w), and the estimated loglikelihood.
localised.structure	NULL, for compatability with the other models.
formula	The formula for the covariate and offset part of the model.
model	A text string describing the model fit.
X	The design matrix of covariates.

Author(s)

Duncan Lee

References

Leroux, B., X. Lei, and N. Breslow (2000). Estimation of disease rates in small areas: A new mixed model for spatial dependence, Chapter Statistical Models in Epidemiology, the Environment and Clinical Trials, Halloran, M and Berry, D (eds), pp. 135-178. Springer-Verlag, New York.

Rushworth, A., D. Lee, and R. Mitchell (2014). A spatio-temporal model for estimating the long-term effects of air pollution on respiratory hospital admissions in Greater London. Spatial and Spatio-temporal Epidemiology 10, 29-38.

Examples

```
###################################################
#### Run the model on simulated data on a lattice
###################################################
#### set up the regular lattice
x.easting <- 1:10
x.northing <- 1:10
Grid <- expand.grid(x.easting, x.northing)
```

```
K <- nrow(Grid)
N <- 10
N.all <- N * K


#### set up spatial neighbourhood matrix W
distance <- as.matrix(dist(Grid))
W <-array(0, c(K,K))
W[distance==1] <-1


#### Simulate the elements in the linear predictor and the data
gamma <- rnorm(n=N.all, mean=0, sd=0.001)
x <- rnorm(n=N.all, mean=0, sd=1)
beta <- 0.1

Q.W <- 0.99 * (diag(apply(W, 2, sum)) - W) + 0.01 * diag(rep(1,K))
Q.W.inv <- solve(Q.W)
phi.temp <- mvrnorm(n=1, mu=rep(0,K), Sigma=(0.1 * Q.W.inv))
phi <- phi.temp
    for(i in 2:N)
    {
    phi.temp2 <- mvrnorm(n=1, mu=(0.8 * phi.temp), Sigma=(0.1 * Q.W.inv))
    phi.temp <- phi.temp2
    phi <- c(phi, phi.temp)
    }

LP <- 3 + x * beta  + phi
mean <- exp(LP)
Y <- rpois(n=N.all, lambda=mean)


#### Run the model
## Not run: model <- ST.CARar(formula=Y~x, family="poisson",  W=W, burnin=10000,
    n.sample=50000, fix.rho.T=TRUE, rho.T=0)
## End(Not run)


#### Toy example for checking
model <- ST.CARar(formula=Y~x, family="poisson",  W=W, burnin=10,
    n.sample=50, fix.rho.T=TRUE, rho.T=0)
```

ST.CARlinear   *Fit a spatio-temporal generalised linear mixed model to data, where*
                                the spatial units have linear time trends with spatially varying inter-
                                cepts and slopes.

Description

Fit a spatio-temporal generalised linear mixed model to areal unit data, where the response variable
can be binomial, Gaussian or Poisson. The linear predictor is modelled by known covariates and an

area specific linear time trend. The intercepts and slopes are spatially autocorrelated and modelled by the conditional autoregressive (CAR) prior proposed by Leroux et al. (2000). The model is similar to that proposed by Bernardinelli et al. (1995) and further details are given in the vignette accompanying this package. Missing values are allowed in the response in this model. Inference is conducted in a Bayesian setting using Markov chain Monte Carlo (MCMC) simulation.

Usage

```
ST.CARlinear(formula, family, data=NULL,  trials=NULL, W, burnin, n.sample, thin=1,
 prior.mean.beta=NULL, prior.var.beta=NULL, prior.mean.alpha=NULL,
 prior.var.alpha=NULL, prior.nu2=NULL, prior.tau2=NULL, fix.rho.slo=FALSE,
 rho.slo=NULL, fix.rho.int=FALSE, rho.int=NULL, MALA=TRUE, verbose=TRUE)
```

Arguments

formula	A formula for the covariate part of the model using the syntax of the lm() function. Offsets can be included here using the offset() function. The response and each covariate should be vectors of length (KN)*1, where K is the number of spatial units and N is the number of time periods. Each vector should be ordered so that the first K data points are the set of all K spatial locations at time 1, the next K are the set of spatial locations for time 2 and so on. Missing values are allowed in the response.
family	One of either 'binomial', 'gaussian' or 'poisson', which respectively specify a binomial likelihood model with a logistic link function, a Gaussian likelihood model with an identity link function, or a Poisson likelihood model with a log link function.
data	An optional data.frame containing the variables in the formula.
trials	A vector the same length and in the same order as the response containing the total number of trials for each area and time period. Only used if family='binomial'.
W	A non-negative K by K neighbourhood matrix (where K is the number of spatial units). Typically a binary specification is used, where the jkth element equals one if areas (j, k) are spatially close (e.g. share a common border) and is zero otherwise. The matrix can be non-binary, but each row must contain at least one non-zero entry. Note, if fix.rho.int=TRUE and rho.int=0 and fix.rho.slo=TRUE and rho.slo=0 then the W matrix is not actually used by the model so a K by K matrix of zeros could be entered instead if a real W matrix cannot be obtained.
burnin	The number of MCMC samples to discard as the burn-in period.
n.sample	The number of MCMC samples to generate.
thin	The level of thinning to apply to the MCMC samples to reduce their temporal autocorrelation. Defaults to 1 (no thinning).
prior.mean.beta	
	A vector of prior means for the regression parameters beta (Gaussian priors are assumed). Defaults to a vector of zeros.
prior.var.beta	A vector of prior variances for the regression parameters beta (Gaussian priors are assumed). Defaults to a vector with values 1000.

prior.mean.alpha

        The prior mean for the average slope of the linear time trend alpha (a Gaussian prior is assumed). Defaults to zero.

prior.var.alpha

        The prior variance for the average slope of the linear time trend alpha (a Gaussian prior is assumed). Defaults to 1000.

prior.nu2   The prior shape and scale in the form of c(shape, scale) for an Inverse-Gamma(shape, scale) prior for the Gaussian error variance nu2. Defaults to c(1, 0.01) and only used if family='Gaussian'.

prior.tau2   The prior shape and scale in the form of c(shape, scale) for an Inverse-Gamma(shape, scale) prior for the random effect variances tau2. Defaults to c(1, 0.01).

fix.rho.slo   Logical, should the spatial dependence parameter for the slope of the linear time trend rho.slo be fixed when fitting the model.

rho.slo   If fix.rho.slo=TRUE, then the value in the interval [0,1] that rho.slo should be fixed to. Not needed if fix.rho.slo=FALSE.

fix.rho.int   Logical, should the spatial dependence parameter for the intercept surface rho.int be fixed when fitting the model.

rho.int   If fix.rho.int=TRUE, then the value in the interval [0,1] that rho.int should be fixed to. Not needed if fix.rho.int=FALSE.

MALA   Logical, should the function use MALA (TRUE, default) or simple random walk (FALSE) updates for the random effects. Not applicable if family='gaussian'.

verbose   Logical, should the function update the user on its progress.

Value

summary.results

        A summary table of the parameters.

samples   A list containing the MCMC samples from the model.

fitted.values A vector of fitted values for each area and time period.

residuals   A matrix with 3 columns where each column is a type of residuals for each area and time period. The types are "response" (raw), "pearson", and "deviance".

modelfit   Model fit criteria including the Deviance Information Criterion (DIC) and its corresponding estimated effective number of parameters (p.d), the Log Marginal Predictive Likelihood (LMPL), the Watanabe-Akaike Information Criterion (WAIC) and its corresponding estimated number of effective parameters (p.w), and the estimated loglikelihood.

localised.structure

        NULL, for compatability with the other models.

formula   The formula for the covariate and offset part of the model.

model   A text string describing the model fit.

X   The design matrix of covariates.

Author(s)

Duncan Lee

References

Bernardinelli, L., D. Clayton, C.Pascuto, C.Montomoli, M.Ghislandi, and M. Songini (1995). Bayesian analysis of space-time variation in disease risk. Statistics in Medicine, 14, 2433-2443.

Leroux, B., X. Lei, and N. Breslow (2000). Estimation of disease rates in small areas: A new mixed model for spatial dependence, Chapter Statistical Models in Epidemiology, the Environment and Clinical Trials, Halloran, M and Berry, D (eds), pp. 135-178. Springer-Verlag, New York.

Examples

```
###################################################
#### Run the model on simulated data on a lattice
###################################################
#### set up the regular lattice
x.easting <- 1:10
x.northing <- 1:10
Grid <- expand.grid(x.easting, x.northing)
K <- nrow(Grid)
N <- 10
N.all <- K * N


#### set up spatial neighbourhood matrix W
distance <- as.matrix(dist(Grid))
W <-array(0, c(K,K))
W[distance==1] <-1


#### Simulate the elements in the linear predictor and the data
x <- rnorm(n=N.all, mean=0, sd=1)
beta <- 0.1
Q.W <- 0.99 * (diag(apply(W, 2, sum)) - W) + 0.01 * diag(rep(1,K))
Q.W.inv <- solve(Q.W)
phi <- mvrnorm(n=1, mu=rep(0,K), Sigma=(0.1 * Q.W.inv))
delta <- mvrnorm(n=1, mu=rep(0,K), Sigma=(0.1 * Q.W.inv))
trend <- array(NA, c(K, N))
time <-(1:N - mean(1:N))/N
    for(i in 1:K)
    {
    trend[i, ] <- phi[i] + delta[i] * time
    }
trend.vec <- as.numeric(trend)
LP <- 4 + x * beta + trend.vec
mean <- exp(LP)
Y <- rpois(n=N.all, lambda=mean)


#### Run the model
## Not run: model <- ST.CARlinear(formula=Y~x, family="poisson", W=W, burnin=10000,
n.sample=50000)
## End(Not run)
```

```
#### Toy example for checking
model <- ST.CARlinear(formula=Y~x, family="poisson", W=W, burnin=10,
n.sample=50)
```

ST.CARlocalised   *Fit a spatio-temporal generalised linear mixed model to data, with a*
                                spatio-temporal autoregressive process and a piecewise constant in-
                                tercept term.

Description

Fit a spatio-temporal generalised linear mixed model to areal unit data, where the response vari-
able can be binomial or Poisson. The linear predictor is modelled by known covariates, a vector of
random effects and a piecewise constant intercept process. The random effects follow the multivari-
ate first order autoregressive time series process proposed by Rushworth et al.(2014), which is the
same as that used in the ST.CARar() function. The piecewise constant intercept component allows
neighbouring areal units to have very different values if they are assigned to a different intercept
component. This model allows for localised smoothness as some pairs of neighbouring areas or time
periods can have similar values while other neighbouring pairs have different values. Furter details
are given in Lee and Lawson (2016) and in the vignette accompanying this package. Inference is
conducted in a Bayesian setting using Markov chain Monte Carlo (MCMC) simulation.

Usage

```
ST.CARlocalised(formula, family, data=NULL, G, trials=NULL, W, burnin, n.sample,
 thin=1, prior.mean.beta=NULL, prior.var.beta=NULL, prior.delta=NULL,
 prior.tau2=NULL, MALA=TRUE, verbose=TRUE)
```

Arguments

formula	A formula for the covariate part of the model using the syntax of the lm() func-tion. Offsets can be included here using the offset() function. The response and each covariate should be vectors of length (KN)*1, where K is the number of spatial units and N is the number of time periods. Each vector should be ordered so that the first K data points are the set of all K spatial locations at time 1, the next K are the set of spatial locations for time 2 and so on.
family	One of either 'binomial' or 'poisson', which respectively specify a binomial likelihood model with a logistic link function, or a Poisson likelihood model with a log link function.
data	An optional data.frame containing the variables in the formula.
G	The maximum number of distinct intercept terms (clusters) to allow in the model.
trials	A vector the same length and in the same order as the response containing the total number of trials for each area and time period. Only used if fam-ily='binomial'.

W	A non-negative K by K neighbourhood matrix (where K is the number of spatial units). Typically a binary specification is used, where the jkth element equals one if areas (j, k) are spatially close (e.g. share a common border) and is zero otherwise. The matrix can be non-binary, but each row must contain at least one non-zero entry.
burnin	The number of MCMC samples to discard as the burn-in period.
n.sample	The number of MCMC samples to generate.
thin	The level of thinning to apply to the MCMC samples to reduce their temporal autocorrelation. Defaults to 1 (no thinning).
prior.mean.beta	
	A vector of prior means for the regression parameters beta (Gaussian priors are assumed). Defaults to a vector of zeros.
prior.var.beta	A vector of prior variances for the regression parameters beta (Gaussian priors are assumed). Defaults to a vector with values 1000.
prior.delta	The prior maximum for the cluster smoothing parameter delta. Defaults to 10.
prior.tau2	The prior shape and scale in the form of c(shape, scale) for an Inverse-Gamma(shape, scale) prior for the random effect variance tau2. Defaults to c(1, 0.01).
MALA	Logical, should the function use MALA (TRUE, default) or simple random walk (FALSE) updates for the random effects.
verbose	Logical, should the function update the user on its progress.

Value

summary.results	
	A summary table of the parameters.
samples	A list containing the MCMC samples from the model.
fitted.values	A vector of fitted values for each area and time period.
residuals	A matrix with 3 columns where each column is a type of residuals for each area and time period. The types are "response" (raw), "pearson", and "deviance".
modelfit	Model fit criteria including the Deviance Information Criterion (DIC) and its corresponding estimated effective number of parameters (p.d), the Log Marginal Predictive Likelihood (LMPL), the Watanabe-Akaike Information Criterion (WAIC) and its corresponding estimated number of effective parameters (p.w), and the estimated loglikelihood.
accept	The acceptance probabilities for the parameters.
localised.structure	
	A vector giving the posterior median of which intercept component (group) each data point is in.
formula	The formula for the covariate and offset part of the model.
model	A text string describing the model fit.
X	The design matrix of covariates.

Author(s)

Duncan Lee

References

Lee, D and Lawson, C (2016). Quantifying the spatial inequality and temporal trends in maternal smoking rates in Glasgow, Annals of Applied Statistics, 10, 1427-1446.

Examples

```
##################################################
#### Run the model on simulated data on a lattice
##################################################
#### set up the regular lattice
x.easting <- 1:10
x.northing <- 1:10
Grid <- expand.grid(x.easting, x.northing)
K <- nrow(Grid)
N <- 10
N.all <- N * K


#### set up spatial neighbourhood matrix W
distance <- as.matrix(dist(Grid))
W <-array(0, c(K,K))
W[distance==1] <-1


#### Simulate the elements in the linear predictor and the data
Q.W <- 0.99 * (diag(apply(W, 2, sum)) - W) + 0.01 * diag(rep(1,K))
Q.W.inv <- solve(Q.W)
phi.temp <- mvrnorm(n=1, mu=rep(0,K), Sigma=(0.1 * Q.W.inv))
phi <- phi.temp
    for(i in 2:N)
    {
    phi.temp2 <- mvrnorm(n=1, mu=(0.8 * phi.temp), Sigma=(0.1 * Q.W.inv))
    phi.temp <- phi.temp2
    phi <- c(phi, phi.temp)
    }
jump <- rep(c(rep(2, 70), rep(4, 30)),N)
LP <- jump + phi
fitted <- exp(LP)
Y <- rpois(n=N.all, lambda=fitted)


#### Run the model
## Not run: model <- ST.CARlocalised(formula=Y~1, family="poisson", G=3, W=W, burnin=10000,
n.sample=50000)
## End(Not run)


#### Toy example for checking
model <- ST.CARlocalised(formula=Y~1, family="poisson", G=3, W=W, burnin=10,
n.sample=50)
```

ST.CARsepspatial	*Fit a spatio-temporal generalised linear mixed model to data, with a common temporal main effect and separate spatial surfaces with individual variances.*

Description

Fit a spatio-temporal generalised linear mixed model to areal unit data, where the response variable can be binomial or Poisson. The linear predictor is modelled by known covariates and two sets of random effects. These include a common temporal main effect and separate time period specific spatial effects with a common spatial dependence parameter but separate variance parameters. Each component is modelled by the conditional autoregressive (CAR) prior proposed by Leroux et al. (2000). Further details are given in Napier et al. (2016) and in the vignette accompanying this package. Inference is conducted in a Bayesian setting using Markov chain Monte Carlo (MCMC) simulation.

Usage

```
ST.CARsepspatial(formula, family, data=NULL, trials=NULL, W, burnin, n.sample,
thin=1, prior.mean.beta=NULL, prior.var.beta=NULL, prior.tau2=NULL, fix.rho.S=FALSE,
 rho.S=NULL, fix.rho.T=FALSE, rho.T=NULL, MALA=TRUE, verbose=TRUE)
```

Arguments

formula	A formula for the covariate part of the model using the syntax of the lm() function. Offsets can be included here using the offset() function. The response and each covariate should be vectors of length (KN)*1, where K is the number of spatial units and N is the number of time periods. Each vector should be ordered so that the first K data points are the set of all K spatial locations at time 1, the next K are the set of spatial locations for time 2 and so on.
family	One of either 'binomial' or 'poisson', which respectively specify a binomial likelihood model with a logistic link function or a Poisson likelihood model with a log link function.
data	An optional data.frame containing the variables in the formula.
trials	A vector the same length and in the same order as the response containing the total number of trials for each area and time period. Only used if family='binomial'.
W	A non-negative K by K neighbourhood matrix (where K is the number of spatial units). Typically a binary specification is used, where the jkth element equals one if areas (j, k) are spatially close (e.g. share a common border) and is zero otherwise. The matrix can be non-binary, but each row must contain at least one non-zero entry. Note, if fix.rho.S=TRUE and rho.S=0 then the W matrix is not actually used by the model so a K by K matrix of zeros could be entered instead if a real W matrix cannot be obtained.
burnin	The number of MCMC samples to discard as the burn-in period.

n.sample	The number of MCMC samples to generate.
thin	The level of thinning to apply to the MCMC samples to reduce their temporal autocorrelation. Defaults to 1 (no thinning).
prior.mean.beta	
	A vector of prior means for the regression parameters beta (Gaussian priors are assumed). Defaults to a vector of zeros.
prior.var.beta	A vector of prior variances for the regression parameters beta (Gaussian priors are assumed). Defaults to a vector with values 1000.
prior.tau2	The prior shape and scale in the form of c(shape, scale) for an Inverse-Gamma(shape, scale) prior for the random effect variances tau2. Defaults to c(1, 0.01).
fix.rho.S	Logical, should the spatial dependence parameter rho.S be fixed when fitting the model.
rho.S	If fix.rho.S=TRUE, then the value in the interval [0,1] that rho.S should be fixed to. Not needed if fix.rho.S=FALSE.
fix.rho.T	Logical, should the temporal dependence parameter rho.T be fixed when fitting the model.
rho.T	If fix.rho.T=TRUE, then the value in the interval [0,1] that rho.T should be fixed to. Not needed if fix.rho.T=FALSE.
MALA	Logical, should the function use MALA (TRUE, default) or simple random walk (FALSE) updates for the random effects.
verbose	Logical, should the function update the user on its progress.

Value

summary.results	
	A summary table of the parameters.
samples	A list containing the MCMC samples from the model. The 'tau2' element of this list has columns (tau2.phi, tau2.delta, tau2.gamma) (the latter if interaction=TRUE). Similarly, the 'rho' element of this list has columns (rho.phi, rho.delta).
fitted.values	A vector of fitted values for each area and time period.
residuals	A matrix with 3 columns where each column is a type of residuals for each area and time period. The types are "response" (raw), "pearson", and "deviance".
modelfit	Model fit criteria including the Deviance Information Criterion (DIC) and its corresponding estimated effective number of parameters (p.d), the Log Marginal Predictive Likelihood (LMPL), the Watanabe-Akaike Information Criterion (WAIC) and its corresponding estimated number of effective parameters (p.w), and the estimated loglikelihood.
accept	The acceptance probabilities for the parameters.
localised.structure	
	NULL, for compatability with the other models.
formula	The formula for the covariate and offset part of the model.
model	A text string describing the model fit.
X	The design matrix of covariates.

Author(s)

Gary Napier

References

Leroux, B., X. Lei, and N. Breslow (2000). Estimation of disease rates in small areas: A new mixed model for spatial dependence, Chapter Statistical Models in Epidemiology, the Environment and Clinical Trials, Halloran, M and Berry, D (eds), pp. 135-178. Springer-Verlag, New York.

Napier, G, D. Lee, C. Robertson, A. Lawson, and K. Pollock (2016). A model to estimate the impact of changes in MMR vaccination uptake on inequalities in measles susceptibility in Scotland, Statistical Methods in Medical Research, 25, 1185-1200.

Examples

```
###################################################
#### Run the model on simulated data on a lattice
###################################################
#### set up the regular lattice
x.easting <- 1:10
x.northing <- 1:10
Grid <- expand.grid(x.easting, x.northing)
K <- nrow(Grid)
N <- 5
N.all <- N * K


#### set up spatial neighbourhood matrix W
distance <- as.matrix(dist(Grid))
W <-array(0, c(K,K))
W[distance==1] <-1


#### Create the spatial covariance matrix
Q.W <- 0.99 * (diag(apply(W, 2, sum)) - W) + 0.01 * diag(rep(1,K))
Q.W.inv <- solve(Q.W)


#### Simulate the elements in the linear predictor and the data
x <- rnorm(n=N.all, mean=0, sd=1)
beta <- 0.1

phi1 <- mvrnorm(n=1, mu=rep(0,K), Sigma=(0.01 * Q.W.inv))
phi2 <- mvrnorm(n=1, mu=rep(0,K), Sigma=(0.01 * Q.W.inv))
phi3 <- mvrnorm(n=1, mu=rep(0,K), Sigma=(0.01 * Q.W.inv))
phi4 <- mvrnorm(n=1, mu=rep(0,K), Sigma=(0.05 * Q.W.inv))
phi5 <- mvrnorm(n=1, mu=rep(0,K), Sigma=(0.05 * Q.W.inv))

delta <- c(0, 0.5, 0, 0.5, 0)
phi.long <- c(phi1, phi2, phi3, phi4, phi5)
delta.long <- kronecker(delta, rep(1,K))
LP <- 4 +  x * beta + phi.long +  delta.long
```

```
mean <- exp(LP)
Y <- rpois(n=N.all, lambda=mean)


#### Run the model
## Not run: model <- ST.CARsepspatial(formula=Y~x, family="poisson", W=W, burnin=10000,
n.sample=50000)
## End(Not run)


#### Toy example for checking
model <- ST.CARsepspatial(formula=Y~x, family="poisson", W=W, burnin=10,
n.sample=50)
```

Index